DOCUMENT RESUME

ED 307 296 TM 013 302

AUTHOR Jones, Gail
TITLE Some Examples of Invariance Procedures in Discriminant Analysis.
PUB DATE Jan 89
NOTE 29p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Houston, TX, January 27, 1989).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150) -- Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Classification; *Discriminant Analysis; Literature Reviews; Multivariate Analysis; Research Methodology; *Statistical Distributions
IDENTIFIERS *Fisher Discriminant Function; *Invariance Principle; Rotations (Factor Analysis)

ABSTRACT

A brief historical background of discriminant analysis is given, with a description of the variety of roles that discriminant analysis can perform. Focus is on the classification role of discriminant analysis and how it can be performed by using Fisher's classification functions or the canonical discriminant functions. A small hypothetical data set consisting of two samples of 20 cases each and for whom the actual classifications (three groups) are known is used to illustrate these methods. The importance of measuring the effectiveness of the classification results by the use of invariance procedures is discussed. The first method of invariance is that of splitting the sample into two halves, developing the functions on one half and then "cross-validating" it on the other half of the sample. The second method of invariance is a comparison of the results obtained by Fisher's classification functions with those obtained by the canonical discriminant functions. The third method of invariance is performed by developing new discriminant function coefficients for the second sample and then rotating the two sets of coefficients from the two samples to "best fit." Contrary to statistical significance, which is often achieved only because a large sample is used and which is often misconstrued as an index of reproducibility, a successful invariance analysis enables the researcher to have confidence that the results will be stable and replicable across samples. Ten data tables are included. (Author/TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED307296

SOME EXAMPLES OF INVARIANCE PROCEDURES IN DISCRIMINANT ANALYSIS

Gail Jones

University of New Orleans

Delgado Community College

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

GAIL JONES

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, January 27, 1989.

TM013302

BEST COPY AVAILABLE

ABSTRACT

A brief historical background of discriminant analysis is given, with a description of the variety of roles that discriminant analysis can perform. This paper focuses on the classification role of discriminant analysis and illustrates how it can be performed by using Fisher's classification functions or the canonical discriminant functions. A small hypothetical data set consisting of two samples of twenty cases each and for whom the actual classifications (three groups) are known is used to illustrate these methods. The importance of measuring the effectiveness of the classification results by the use of invariance procedures is discussed. The first method of invariance is that of splitting the sample into two halves, developing the functions on one half and then "cross-validating" it on the other half of the sample. The second method of invariance is a comparison of the results obtained by Fisher's classification functions with those obtained by the canonical discriminant functions. The third method of invariance is performed by developing new discriminant function coefficients for the second sample and then rotating the two sets of coefficients from the two samples to "best fit."

As developed by Fisher in 1936, the basic purpose of discriminant analysis was to provide a means of classifying an object into one of two categories. Rao extended the number of categories to more than two in 1948, but the basic purpose remained the same. Since the 1960's, however, the utilization of discriminant analysis has increased greatly. Huberty (1975) describes the roles of discriminant analysis as classification, separation, discrimination, and estimation. These new roles have extended the use of discriminant analysis to one of analyzing group differences. Separation, as defined by Huberty (1975), refers to significance testing via MANOVA; discrimination is the interpretation of the linear discriminant functions and their coefficients; estimation refers to characterizing intergroup differences and strength of relationship; and classification is the development of rules for assigning individuals to groups. According to Cooley and Lohnes (1971), the discriminant model is a kind of factor analysis which extracts orthogonal factors for the purpose of displaying and capitalizing upon differences among criterion groups.

The variety of terms with which the activities of discriminant analysis are described can be very confusing to the novice researcher. Generally speaking, discriminant analysis involves one of two basic activities: interpreting group differences and classifying cases into groups (Klecka, 1980) or, even more simply expressed, description and prediction, respectively (Afifi & Clark, 1984). The most logical

categorization of these activities comes from Hand (1981), who describes *discrimination* as the process of developing classification rules from samples of classified objects and then uses the term *classification* to refer to applying these rules to new objects whose classes are not known. Since the statisical analysis for the examples in the present paper were performed with the statistical package SPSS-X, it should be noted that Norusis (1985) describes discriminant analysis as a procedure for predicting group membership for new cases whose group membership is undetermined. The output given, however, permits the researcher to engage in all four of the activities described by Huberty (1975).

Discriminant analysis, like all of the parametric univariate and multivariate techniques, is based upon the linear model. The first step of discriminant analysis is to find optimal linear discriminant functions which will best discriminate between the groups and minimize the probability of misclassification. Most researchers agree on two basic assumptions that must be met: (a) each group must be taken from a population which has a multivariate normal distribution, and (b) the population covariance matrices for each group must be equal. Klecka (1980) adds a third assumption that no discriminating variable can be a linear combination of other discriminating variables or be perfectly correlated with another discriminating variable. However, this first assumption is based on testing the statistical significance of the resulting discriminant

functions or variables in order to discard those functions or variables which do not contribute to group separation. If the researcher intends to employ all of the functions and variables in the analysis, then no tests of statistical significance will be employed and the first basic assumption need not necessarily be met.

Thompson (1987, 1988) emphasizes that statistical significance does not inform us about the importance of results and stresses that the researcher should look for other analogs which indicate the noteworthiness of results. Fish (1986) states that the most interesting research results are those which, however significant statistically, can be generalized from a sample to a larger population. Carver (1978) argues that the confounding of statistical significance and reproducibility is one of the major prevailing misconceptions about statistical significance testing.

In discriminant analysis, the analog which indicates the effectiveness of the results is invariance analysis. Invariance procedures are used to determine how stable the results will be across different samples. Because invariance analysis is still a relatively new field, several exploratory methods of invariance are discussed, along with one method which is widely accepted. The present paper focuses on the classification role of discriminant analysis and then provides the results of three different invariance procedures. A concrete example is discussed.

Fisher's Classification Functions

The activity of classification can be performed by using either the discriminating variables or the discriminant functions to predict the group to which a case most likely belongs. In his original classification, Fisher used the theory of maximizing the ratio of between-group differences to within-groups differences to derive classification functions (Klecka, 1980). He defined the separation between two groups in a particular direction as the distance between the means of the two groups standardized for the within-group variance in the specified direction (Hand, 1981).

No distributional assumptions are required for the maximization of this ratio and the linear classification functions which result (Hand, 1981; Huberty, 1975). Consequently, no tests are made for the significance of the classification or the dimensionality of a discriminant function space (Klecka, 1980).

The classification function, as developed by Fisher, has the following form:

$$c_k = b_{ko} + b_{k1}X_1 + b_{k2}X_2 + \ldots + b_{kp}X_p$$

where c_k is the score for group k and the b's are coefficients that need to be derived. Each case is classified into the group with the highest score or largest c. The derivation of these coefficients is mathematically complex and beyond the scope of this paper. These classification function coefficients (b's) are not interpreted because

they are not standardized and there is a different function for each group. Klecka (1980) states that "the scores also lack intrinsic value because they are arbitrary numbers which have the property that the case resembles most closely that group on which it has the highest score" (p. 43).

Before looking at the classification function coefficients, the research situation for the concrete example must be described. Prediction into one of three graduate programs is desired: Masters plus 30, Ed.D., or Ph.D. This prediction is to be based on the profiles of 20 graduate students who have entered these programs and been successful. These 20 students constitute sample one, the sample on which the classification function is to be based. Another 20 students for whom prediction is desired, but whose group membership is already known, will constitute sample two. The first 20 cases are identified as sample one and the second 20 cases as sample two for convenience, although ideally the cases should be randomly assigned. The discriminating variables are NTE percentiles, undergraduate gpa (based on a 4.0 scale), and GRE scores.

Table 1 presents the classification function coefficients for sample one. These coefficients are not used for classification purposes in the SPSS-X statistical package. Consequently, it was necessary to use the six transformation statements in Table 2 to develop the classification scores.

INSERT TABLES 1 AND 2 ABOUT HERE.

Since there are three groups, it was necessary to compute three linear combinations for each case in the target sample. A case is classified into the group that produces the highest classification score. Again, since the only output provided by SPSS-X is Fisher's classification function coefficients, the researcher must evaluate the classification scores to display the classification results and to evaluate the misclassification rate. This information is provided in Table 3 and will later be compared with the results obtained by using the discriminant function coefficients.

INSERT TABLE 3 ABOUT HERE.

Canonical Discriminant Functions

When the number of variables becomes quite large, the use of discriminant functions can greatly reduce the number of computations since the number of functions can be no greater than the smaller of the number of groups minus one or the number of variables. Classification performance of classification analysis can also deteriorate as the number of variables increases. Under such circumstances, increased reliability will result if the dimensional space is reduced by using the discriminant functions rather than the variables (Hand, 1981). Historically, discriminant analysis also was useful in the era before widespread availability of computers, since

the analysis may be computationally easier than classification analysis.

Fisher was the first to propose the discriminant function as a solution to the problem of using information from a number of correlated variables to classify objects (Tatsuoka & Tiedeman, 1954). Once the linear combinations are developed, there are several classification procedures, both parametric and nonparametric, that can be used to decide in which group a case belongs. Most of these procedures are based on a comparison of a case's position to each of the group centroids in order to locate the closest one. A centroid is an average function score of the objects in a given class.

Initially, the use of Fisher's linear discriminant function for classification purposes for the two-group situation was based on simple Euclidean distance (Klecka, 1980). Several years after Fisher's work was introduced, Welch recognized the relationship of posterior probability of group membership to the linear discriminant function when he proved that the assignment procedure based on the linear discriminant function minimizes the probability of misclassification under certain restrictions. Today, this closeness is generally measured by a "probabalistic" notion of distance, rather than a Euclidean distance.

Today, a generalization of Fisher's approach, referred to as canonical variate analysis, is the discriminant analysis method most frequently used in computer packages (Hand, 1981). When the

canonical discriminant functions are used, the information contained in the predictor variables is summarized in a single index, the discriminant score (Norusis, 1985). The canonical discriminant function is expressed as follows:

$$f_{km} = u_0 + u_1X_{1km} + u_2X2_{km} + \ldots + u_pX_{pkm}$$

where f_{km} is the discriminant score for case m in group k, the X's are the values of the independent variables, and the u's are coefficients estimated from the data. The u's are chosen so that the group means (centroids) of each set of discriminant function scores differ as much as possible between the groups, or that for the discriminant scores, the ratio of between-groups sum-of-squares to within-groups sum-of-squares is maximized. As each successive function is derived, the values on that function are not correlated with the values on the preceding functions (Klecka, 1980). In other words, discriminant functions are orthogonal or uncorrelated, just like factors in factor analysis.

Once the discriminant scores have been obtained, SPSS-X uses a classification rule based on Bayes' Rule (Norusis, 1985):

$$P(G_i/D) = \frac{P(D/G_i)P(G_i)}{\sum P(D/G_i)P(G_i)}$$

where $P(G_i)$, the prior probability, is an estimate of the likelihood that a case belongs to a particular group (i) when no information about it is available; $P(D/G_i)$, the conditional probability, is the calculated probability that a specific discriminant score will be obtained when membership in the group is estimated; and $P(G_i/D)$, the posterior probability, is the probability of membership in a specific group given that a specific discriminant score was obtained. Since it is assumed that the group membership is not known, the posterior probability of membership in each group for each case will be calculated according to Bayes' Rule. Each case is then classified into the group for which it has the largest posterior probability. Table 4 gives the results of the classifications by the discriminant function method for the same 20 subjects for whom classification function scores were presented in Table 3. Tables 5 and 6 present the standardized discriminant function and structure coefficients.

INSERT TABLES 4, 5, AND 6 ABOUT HERE.

The present concern with discriminant functions has been solely with the effectiveness of the classification role. However, the SPSS-X output provides a variety of other information which enables the researcher to judge the importance of the discriminant functions. The relative importance of each function is expressed as a relative percentage; the canonical correlations indicate how closely the

function is related to the predictor variables; and Wilks' lambda reveals how much discriminating power in the original predictor variables has not been removed by the discriminant functions. The standardized function coefficients indicate the relative contribution of each variable to each function, while the structure coefficients indicate the correlations between the variables and the functions (Klecka, 1975).

A Widely Accepted Invariance Procedure

Klecka (1980) states that the proportion of cases correctly classified indicates the accuracy of the procedure and indirectly confirms the degree of group separation. However, both Klecka (1980) and Norusis (1985) state that, since the model fits the sample from which it is derived, this classification rate is an inflated estimate of true performance in the population and yields only an apparent misclassification rate. They indicate that, with large enough sample sizes, a preferred estimate of the true misclassification rate would be to split the sample into two parts, use one part to derive the discriminant function, and the other part to test it. The error rate in the test sample is a better measure of the function's effectiveness. This procedure is known as cross-validation and has been more frequently applied in multiple regression. The degree of correspondence between the actual group membership of the cases in

the second group and the predictions made for these cases on the basis of the discriminant functions developed from the first group indicates how successful the cross-validation has been (Afifi & Clark, 1984; Huck, Cormier, & Bounds, 1974).

Since the membership of each case in both samples one and two is known, the number of correct classifications made by the discriminant functions can be determined for both samples. The cases for which the highest group is the same as the actual group are called hits, while the cases for which the highest group is not the same as the actual group are called misses. The classification results give the number of hits on the diagonal and the number of misses off the diagonal. Table 7 indicates a correct classification rate of 85% for sample one, on which the discriminant functions were developed, while Table 8 gives a slightly lower correct classification rate of 80% for sample two, the test sample for which cases were classified using the functions derived using sample one data. Thus, there is an apparent misclassification rate of 15% and a true misclassification rate of 20%. The discriminant functions appear to be reasonably invariant across samples.

INSERT TABLES 7 AND 8 ABOUT HERE.

The percentage of correctly classified cases must also be evaluated against the number of cases which would correctly be classified by chance alone. Since there are three groups, assuming

equal group size, a correct classification rate of 33% could be expected by chance alone. A comparison of the "true" classification rate of 80% with an expected chance classification rate of 33% indicates that the discriminant functions are performing quite well indeed.

A Comparison of Two Sets of Classification Results

A comparison of the classifications obtained by the classification functions with those obtained by the discriminant functions can now be undertaken. A comparison of the results in Tables 3 and 4 reveals that the classifcation results are similar, but not identical. Using the classification functions yielded a correct classification rate of 90% on sample one, while the discriminant functions obtained a correct classification rate of 85%. According to Norusis (1985) and Klecka (1980), the final classifications of the two procedures will be identical if all canonical discriminant functions are used. Klecka (1980) adds the stipulation that the group covariance matrices must be equal for these two methods to produce identical results. A look at the Box's M statistic reveals that it is statistically significant at the .2421 level, indicating that the null hypothesis that the group covariance matrices are equal can not be rejected at conventional probability levels. However, the covariance matrices are not exactly equal, or the calculated *p*-value for the test would be 1.0. According to Tatsuoka (1971), "when the K population

dispersion matrices are not assumed to be equal ... the identity of classificatory results in the original p-dimensional space and the reduced (K-1) dimensional space no longer holds" (pp. 232-233). However, Tatsuoka (1971) also notes that "classificatory results in the original predictor space and the reduced discriminant space tend to be closely similar even when the separate group dispersion matrices are used, provided they are not drastically different" (p. 236).

Although only two classification rules are presently being considered, Huberty indicates that many classification rules are possible. Five different classification rules are evaluated in a 1974 paper by Huberty, with the conclusion that "it may be well to only consider classification accuracy of rules applied in a cross-validation setting" (p. 81). In his 1975 paper on discriminant analysis, he lists nine different classification statistics that can be used. He indicates that the choice of a method depends on whether the covariance matrices are equal, whether one wishes to use differential prior probabilities, and may even be based on one's preference for use of statistics based on the classical approach or on the Bayesian solution.

Invariance by Rotation of the Standardized Function Coefficients

Like all correlational methods, discriminant analysis tends to capitalize on chance. Thompson (1981) notes that most extraction

procedures reach their criterion by such capitalization. In the case of factor analysis, Gorsuch (1983) recommends that an estimate of the magnitude of this capitalization can be found by randomly splitting the sample in half and comparing the factors extracted from both data sets. Thompson (1981) asserts that this same logic can be generalized to all correlational methods, all of which capitalize on chance.

At this point, one set of discriminant functions has been generated based on the data in sample one. The data in sample two will now be used to generate another set of discriminant functions. The function coefficients obtained from these results are presented in Table 9. The result will be two sets of coefficients (Table 5 with Table 9) which can be compared using "best fit" rotation methods. Thompson (1986) provides a discussion of an empirical method developed by Kaiser, Hunka, and Bianchini for "relating" factors based upon data from different samples. The method consists of projecting the two sets of factors into the same space and calculating the cosines of the angles among the factors across the two solutions. These cosines are, in fact, correlation coefficients and will provide a measure of the relatedness of the two sets of factors. Thompson (1981, 1986) states that these coefficients are analagous to test-retest coefficients and has called them invariance coefficients, although they may also be utilized for substantive interpretations and as adequacy coefficients (Thompson, 1983). The function coefficients

have been rotated instead of the structure coefficients because the primary concern is to investigate the similarity of function equations rather than the similarity of the distributions of variance (Thompson, 1984).

Using the RELATE program by Veldman, the standardized discriminant function coefficients from sample one are input as matrix A and those from sample two as matrix B. The results of the rotation indicate a cosine between the functions of .99. Likewise, the cosines for the three variables are in the same range: .99, .97, and .99, for NTE percentiles, undergraduate gpa, and GRE scores, respectively. According to examples by Kaiser et al. and Gorsuch which were mentioned by Thompson (1986), these cosines fall into the range of reasonable to exceptional. This rotation to "best fit" yielded extremely high invariance coefficients, which in turn indicates that classification based on these discriminant functions may be generalized for future research.

It should be observed, however, that little is known about the characteristics of the various invariance estimates. As a result, Thompson (1984) suggested that the researcher might wish to employ several strategies to obtain both "lower bound" and "upper bound" estimates of the degree of capitalization on sampling specificity. Thompson (1986) used Monte Carlo techniques to generate a partial test distribution for factor cosines, since the minimally acceptable values for these cosines have not been clearly established. The

reader is reminded that the cosine for 0 degrees is I and the cosine for 90 degrees is 0. Therefore, the greater the relationship between the factors, the closer to 1 the cosine will be.

Summary

When the primary goal is to develop a mathematical model which can predict well the classifications of unclassified objects, the most important guide regarding the model's accuracy is the percentage of correct classifications. However, when this percentage reflects the accuracy of the same sample on which the discriminant functions were derived, it is an inflated estimate of the model's performance for the full population. Accordingly, the researcher can increase the reliability of the percentage of correct classifications by using "cross-validation" techniques. Alternative methods of invariance analysis such as those discussed in the present paper may also be explored to validate the accuracy of the model.

Contrary to statistical significance, which is often achieved only because a large sample was used and which is often misconstrued as an index of reproducibility, a successful invariance analysis enables the researcher to have confidence that the results will be stable and replicable across samples. As Thompson (1986) states, "researchers have increasingly recognized the critical nature of replication as the ultimate test of scientific findings and some have argued that replicability should replace significance testing as part of a new logic of truth testing " (p. 27).

It should be noted that, at this time, the interpretation of invariance results is basically a matter of subjective judgment. However, the lack of specific criteria by which to assess the invariance results in no way diminishes the need for performing invariance procedures to evaluate the successful replicability or generalizability of the results.

REFERENCES

Afifi, A. A., & Clark, V. (1984). Computer-aided multivariate analysis. Belmont, CA: Wadsworth.

Cooley, W. W., & Lohnes, P. R. (1971). Multivariate data analysis. New York: Wiley.

Fish, L. (1986, November). The importance of invariance procedures as against tests of statistical significance. Paper presented at the annual meeting of the Mid-South Educational Research Association, Memphis.

Gorsuch, R. L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hand, D. J. (1981). Discrimination and classification. Chichester, England: Wiley.

Huberty, C. J. (1975). Discriminant analysis. Review of Educational Research, 45, 543-598.

Carver, R. C. (1978). The case against statistical significance testing. Harvard Educational Review, 48, 378-399.

Huberty, C. J., & Bloomers, P. J. (1974). An empirical comparison of the accuracy of selected multivariate classification rules. Multivariate Behavioral Research, 9, 59-84.

Huck, S. W., Cormier, W. H., & Bounds, W. G. (1974). Reading statistics and research. New York: Harper & Row.

Klecka, W. R. (1975). Discriminant analysis. In Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (Eds.). Statistical package for the social sciences (2nd ed.) (pp. 434-467). New York: McGraw-Hill.

Klecka, W. R. (1980). Discriminant analysis. Beverly Hills, CA: Sage.

Norusis, M. J. (1985). SPSSX: Advanced statistics guide. Chicago: McGraw-Hill.

Tatsuoka, M. M., & Tiedeman, D. V. (1954). Discriminant analysis. Review of Educational Research, 24, 402-420.

Tatsuoka, M. M. (1971). Multivariate analysis. New York: Wiley.

Thompson, B. (1981). Utility of invariance coefficients. Perceptual and Motor Skills, 52, 708-710.

Thompson, B. (1984). Canonical correlation analysis: Uses and interpretation. Beverly Hill, CA: Sage.

Thompson, B. (1986, April). A partial test distribution for cosines among factors across samples. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Thompson, B. (1987, April). The use (and misuse) of statistical significance testing. Paper presented at the annual meeting of the American Educational Research Association, Washington.

Thompson, B. (1988). A note about significance testing. Measurement and Evaluation in Counseling and Development, 20 (4), 146-147.

Table 1

Classification Function Coefficients for Sample One

CLASSIFICATION FUNCTION COEFFICIENTS
(FISHER'S LINEAR DISCRIMINANT FUNCTIONS)

DEGREE =	1	2	3
NTE	6.153809	6.845623	7.686987
UNDERGPA	90.87343	94.07794	106.5835
GRE	-0.1559736	-0.1695838	-0.2050293
(CONSTANT)	-294.4094	-345.7796	-421.4329

Table 2

Transformation Statements to Obtain Classification Scores

```
COMPUTE GRP1 = 6.153809*NTE + 90.87343*UNDERGPA - 0.1559736*GRE - 294.4094
COMPUTE GRP2 = 6.85623*NTE + 94.07794*UNDERGPA - 0.1695838*GRE - 345.7796
COMPUTE GRP3 = 7.686987*NTE + 106.5835*UNDERGPA - 0.2050293*GRE - 421.4329
IF (GRP1 > MAX(GRP2,GRP3)) CLASS = 1
IF (GRP2 > MAX(GRP1,GRP3)) CLASS = 2
IF (GRP3 > MAX(GRP1,GRP2)) CLASS = 3
```

Table 3

Class Assignments using Fisher's Classification Method

						Classification Scores			
CASEID	SAMPLE	DEGREE	NTE	UNDERGPA	GRE	GRP1	GRP2	GRP3	CLASS
1	1	1	75	2.75	975	264.95	261.81	248.29	1.00
2	1	1	72	3.00	1000	265.31	260.52	246.75	1.00
3	1	3	90	3.45	1100	401.38	409.31	412.58	3.00
4	1	2	82	3.25	1050	341.77	344.12	340.02	2.00
5	1	3	89	3.15	1010	382.00	389.49	391.37	3.00
6	1	2	80	3.20	1090	318.68	318.92	311.11	2.00
7	1	2	89	3.15	1150	360.16	365.75	362.66	2.00
8	1	3	95	3.20	1200	393.83	403.11	403.86	3.00
9	1	3	91	3.55	1175	404.92	412.85	415.54	3.00
10	1	2	87	3.10	1170	340.19	343.94	337.86	2.00
11	1	3	94	3.55	1210	417.92	427.49	431.43	3.00
12	1	3	97	3.85	1300	449.61	461.02	468.01	3.00
13	1	1	80	3.05	1005	318.31	319.22	312.55	2.00
14	1	1	84	3.15	1050	344.99	348.43	344.73	2.00
15	1	3	93	3.40	1190	401.26	409.91	411.86	3.00
16	1	2	85	3.10	1075	342.70	346.34	341.96	2.00
17	1	2	81	3.00	1005	319.92	321.38	314.91	2.00
18	1	1	75	2.90	990	276.25	273.38	261.20	1.00
19	1	2	89	3.25	1175	365.35	370.92	368.20	2.00
20	1	1	77	3.05	1110	283.47	280.85	267.96	1.00
21	2	1	75	3.00	1030	279.09	276.00	263.66	1.00
22	2	1	82	3.15	1045	333.46	335.56	330.38	2.00
23	2	1	81	3.10	1060	320.42	321.46	314.29	2.00
24	2	1	85	3.05	1075	338.16	341.64	336.63	2.00
25	2	1	78	3.15	1020	312.75	312.38	304.76	1.00
26	2	1	79	2.95	1010	302.28	302.11	293.18	1.00
27	2	1	83	3.00	1040	326.76	329.15	323.11	2.00
28	2	3	94	3.85	1225	442.84	453.17	460.33	3.00
29	2	3	89	3.50	1170	388.85	395.28	395.87	3.00
30	2	3	91	3.45	1195	392.71	400.05	400.79	3.00
31	2	3	95	3.35	1190	409.02	418.92	421.90	3.00
32	2	3	98	3.90	1335	454.85	466.64	473.85	3.00
33	2	3	92	3.65	1225	412.36	420.64	423.64	3.00
34	2	3	90	3.40	1165	386.69	393.58	393.92	3.00
35	2	2	81	2.95	1005	315.37	316.67	309.58	2.00
36	2	2	83	3.05	1100	321.95	323.68	316.13	2.00
37	2	2	88	3.15	1175	350.11	354.65	349.85	2.00
38	2	2	88	3.25	1190	356.86	361.52	357.43	2.00
39	2	2	90	3.20	1180	366.18	372.22	369.53	2.00
40	2	2	87	3.25	1135	359.28	363.99	361.02	2.00

Table 4

Predicted Class Assignments (Highest Group) Using Discriminant Function Method

SE NUM	SEL	ACTUAL GROUP		HIGHEST PROBABILITY GROUP	P(D/G)	P(G/D)	2ND HIGHEST GROUP	P(G/D)	DISCRIMINANT SCORES...	
1	YES	1		1	0.4720	0.9810	2	0.0190	-2.9845	-1.7385
2	YES	1		1	0.3004	0.9962	2	0.0038	-3.8850	-0.4461
3	YES	3		3	0.5052	0.9856	2	0.0144	1.6719	2.0623
4	YES	2		2	0.2838	0.7897	1	0.1793	-0.7732	1.0402
5	YES	3		3	0.7739	0.9437	2	0.0562	1.5598	1.2968
6	YES	2	**	1	0.6385	0.6468	2	0.3529	-1.6174	-0.0919
7	YES	2		2	0.4333	0.8873	3	0.1042	0.9386	-1.0406
8	YES	3		3	0.0716	0.8531	2	0.1469	2.6522	-1.2231
9	YES	3		3	0.7510	0.9748	2	0.0252	1.7211	1.5974
10	YES	2		2	0.3497	0.9391	1	0.0555	0.1831	-1.8388
11	YES	3		3	0.9513	0.9929	2	0.0071	2.5173	1.1541
12	YES	3		3	0.3606	0.9997	2	0.0003	3.2639	2.0136
13	YES	1	**	2	0.4253	0.5168	1	0.4817	-1.3374	0.2168
14	YES	1	**	2	0.7045	0.8780	1	0.0689	-0.2036	0.3939
15	YES	3		3	0.7808	0.9494	2	0.0506	2.1918	0.3305
16	YES	2		2	0.9796	0.9126	1	0.0591	-0.0297	-0.3467
17	YES	2		2	0.6616	0.6446	1	0.3530	-1.0526	-0.1064
18	YES	1		1	0.8200	0.9750	2	0.0250	-2.9539	-0.8784
19	YES	2		2	0.5254	0.8483	3	0.1434	0.8924	-0.7177
20	YES	1		1	0.5656	0.9688	2	0.0312	-2.7548	-1.6774
21	NO	1		1	0.7700	0.9799	2	0.0201	-3.0666	-0.8059
22	NO	1	**	2	0.5910	0.7659	1	0.2233	-0.8158	0.3833
23	NO	1	**	2	0.5849	0.5428	1	0.4562	-1.2319	-0.2843
24	NO	1	**	2	0.9530	0.9152	1	0.0696	-0.0621	-0.7169
25	NO	1		1	0.3992	0.7678	2	0.2320	-1.9735	0.6126
26	NO	1		1	0.8289	0.7329	2	0.2670	-1.7415	-0.6541
27	NO	1	**	2	0.9245	0.8152	1	0.1802	-0.5736	-0.5967
28	NO	3		3	0.1027	0.9997	2	0.0003	2.6451	3.1247
29	NO	3		3	0.5082	0.8208	2	0.1784	1.0766	1.2167
30	NO	3		3	0.7072	0.8450	2	0.1547	1.5676	0.5232
31	NO	3		3	0.5271	0.9818	2	0.0182	2.7937	0.0544
32	NO	3		3	0.3360	0.9997	2	0.0003	3.4581	1.8465
33	NO	3		3	0.8246	0.9816	2	0.0184	1.8812	1.5500
34	NO	3		3	0.6231	0.7844	2	0.2150	1.3511	0.6068
35	NO	2		2	0.6806	0.6080	1	0.3908	-1.0850	-0.4766
36	NO	2		2	0.6141	0.7006	1	0.2986	-0.8074	-1.2283
37	NO	2		2	0.4394	0.9543	1	0.0258	0.5105	-1.5050
38	NO	2		2	0.6566	0.9373	3	0.0401	0.5087	-1.0151
39	NO	2		2	0.3130	0.8461	3	0.1487	1.1550	-1.1243
40	NO	2		2	0.7772	0.8677	3	0.1126	0.4356	-0.1439

Table 5

Standardized Discriminant Function Coefficients

Obtained Using Sample One

STANDARDIZED CANONICAL DISCRIMINANT FUNCTION COEFFICIENTS

	FUNC 1	FUNC 2
NTE	1.13777.	0.26533
UNDERGPA	0.56031	-1.12087
GRE	-0.73872	1.00044

Table 6

Structure Coefficients Obtained Using Sample One

STRUCTURE MATRIX:

POOLED WITHIN-GROUPS CORRELATIONS BETWEEN DISCRIMINATING VARIABLES AND CANONICAL DISCRIMINANT FUNCTIONS
(VARIABLES ORDERED BY SIZE OF CORRELATION WITHIN FUNCTION)

	FUNC 1	FUNC 2
NTE	0.86443*	0.48065
UNDERGPA	0.57195*	-0.36420
GRE	0.41154	0.46409*

Table 7

Classification Results for Sample One

CLASSIFICATION RESULTS FOR CASES SELECTED FOR USE IN THE ANALYSIS -

ACTUAL GROUP	NO. OF CASES	PREDICTED GROUP MEMBERSHIP 1	2	3
GROUP 1	6	4 66.7%	2 33.3%	0 0.0%
GROUP 2	7	1 14.3%	6 85.7%	0 0.0%
GROUP 3	7	0 0.0%	0 0.0%	7 100.0%

PERCENT OF "GROUPED" CASES CORRECTLY CLASSIFIED: 85.00%

Table 8

Classification Results for Sample Two

CLASSIFICATION RESULTS FOR CASES NOT SELECTED FOR USE IN THE ANALYSIS -

ACTUAL GROUP	NO. OF CASES	PREDICTED GROUP MEMBERSHIP 1	2	3
GROUP 1	7	3 42.9%	4 57.1%	0 0.0%
GROUP 2	6	0 0.0%	6 100.0%	0 0.0%
GROUP 3	7	0 0.0%	0 0.0%	7 100.0%

PERCENT OF "GROUPED" CASES CORRECTLY CLASSIFIED: 80.00%

Table 9

Standardized Discriminant Function Coefficients

Obtained Using Sample Two

STANDARDIZED CANONICAL DISCRIMINANT FUNCTION COEFFICIENTS

	FUNC 1	FUNC 2
NTE	0.84752	-0.04638
UNDERGPA	0.82419	-1.25826
GRE	-0.56037	1.52413

Table 10

Structure Coefficients Obtained Using Sample Two

STRUCTURE MATRIX:

POOLED WITHIN-GROUPS CORRELATIONS BETWEEN DISCRIMINATING VARIABLES AND CANONICAL DISCRIMINANT FUNCTIONS
(VARIABLES ORDERED BY SIZE OF CORRELATION WITHIN FUNCTION)

	FUNC 1	FUNC 2
UNDERGPA	0.85100*	-0.14727
NTE	0.83363*	0.50684
GRE	0.72793*	0.54996